COMMISSION REGULATION (EC) No 1981/2003

of 21 October 2003

implementing Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning Community statistics on income and living conditions (EU-SILC) as regards the fieldwork aspects and the imputation procedures

(Text with EEA relevance)

THE COMMISSION OF THE EUROPEAN COMMUNITIES,

Having regard to the Treaty establishing the European Community,

Having regard to Regulation (EC) No 1177/2003 of the European Parliament and of the Council of 16 June 2003 on Community statistics on income and living conditions (EU-SILC) [1], and in particular Article 15(2)(e) thereof,

Whereas:

(1) Regulation (EC) No 1177/2003 established a common framework for the systematic production of Community statistics on income and living conditions, encompassing comparable and timely cross-sectional and longitudinal data on income and on the level and composition of poverty and social exclusion at national and European Union levels.

(2) Pursuant to Article 15(2)(e) of Regulation (EC) No 1177/2003, implementing measures are necessary to harmonise methods and definitions concerning the fieldwork aspects and the imputation procedures.

(3) The measures provided for in this Regulation are in accordance with the opinion of the Statistical Programme Committee,

HAS ADOPTED THIS REGULATION:

Article 1

The fieldwork aspects, including household and personal identification numbers, the rules and guidelines for applying substitutions and the imputation procedures, shall be as laid down in the Annex.

Article 2

This Regulation shall enter into force on the 20th day following that of its publication in the *Official Journal of the European Union*.

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Brussels, 21 October 2003.

For the Commission
Pedro SOLBES MIRA
Member of the Commission

[1] OJ L 165, 3.7.2003, p. 1.

A E

1. DEFINITIONS

For the purposes of this Regulation, the following definitions shall apply to the longitudinal component of EU-SILC:

(a) Initial sample: refers to the sample of households or persons at the time it is selected for inclusion in EU-SILC.

(b) Sample persons: means all or a subset of the members of the households in the initial sample who are over a certain age.

(c) Age limit used to define sample persons:

In the case of a four-year panel, this age limit shall not be higher than 14 years. In countries with a four-year panel using a sample of addresses or households, all household members aged 14 and over in the initial sample shall be sample persons. In countries with a four-year panel using a sample of persons, this shall mean selecting at least one such person per household.

The abovementioned minimum age limit shall be lower in the case of a longer panel duration. For a panel duration exceeding eight years, members of all ages in the initial sample shall be sample persons, and children born to sample women during the time the mother is in the panel shall be included as sample persons.

(d) Sample household: means a household containing at least one sample person. A sample household shall be included in EU-SILC for the collection or compilation of detailed information where it contains at least one sample person aged 16 or more.

(e) Co-residents (non-sample persons): are all current residents of a sample household other than those defined above as sample persons.

(f) Current household members: means current members of the household at the time of data collection or compilation. The conditions for classifying a person as a household member are laid down in Annex I to Commission Regulation (EC) No 1980/2003 [1] on definitions and updated definitions.

(g) Former household member: refers to a person who is not a current household member and was not recorded as a household member in that household in the previous wave, but who lived in the household for at least three months during the income reference period.

(h) Entire household:

A sample household is said to be entire (whole) if it remains as one household, without forming an additional household and without the household disappearing, even though there may have been changes in its composition from the previous wave due to deaths, members moving out of scope or co-residents leaving the household, people joining the household or births.

(i) Initial/Split-off household:

Sample household from wave x is said to have been 'split' if its sample persons from wave x reside at the time of wave x+1 in more than one private household within the national territories included in the target population.

When a split has occurred, one (and only one) of the resulting households shall be defined as the 'initial' household, while one or more of the others are termed 'split-off' households.

The following approach shall be used in order to distinguish between 'initial' and 'split-off' households:

— if any sample person of the wave x still lives at the same address as the last wave, then his/her household shall be defined as the 'initial' household. All sample persons who have moved shall form one or more 'split-off' households,

[1] See page 1 of this Official Journal.

— if no sample person lives at the address of the last wave, then the household of the sample person who had the lowest person number in the register for the last wave shall be the initial household. Where this person is no longer alive or in a private household within the national territory of the target population, the initial household shall be the household of the sample person with the lowest person number.

(j) Fusion:

Sample persons from different sample households from the previous wave join together to form a new household.

For the cross-sectional and longitudinal component of EU-SILC, the following definition shall apply:

(k) Person number: refers to the number assigned in the 'household register' to the person the first time he/she is recorded as a household member. In the cross-sectional component, and in new households in the longitudinal component, it should correspond to the person's line position in the 'household register'.

2. IDENTIFICATION NUMBERS FOR PERSONS AND HOUSEHOLDS

2.1. **Household identification number**

1. In the cross-sectional component, the household fixed identification number (ID) is a sequential number.

2. In the longitudinal component, the household identification number (ID) is composed of the household number and the split number. The household number is a sequential number. The split number for the first wave shall always take the value '00'.

Where the household remains entire, it shall keep the household number and split number from one wave to the next.

In the case of a split-off, the initial household shall keep the household number and split number from one wave to the next. The other households, i.e. split-off households, shall keep the same household number, but shall be assigned the next available unique split number in sequence.

In the case of a fusion of two sample households, if the new household is still at a previous address, it shall retain the household number and split number of the household that was at that address in the previous wave. If the new household is at a new address, the household number and split number of the household of the sample person who now has the lowest person number in the 'household register' shall be used.

2.2. **Personal identification number**

1. In the cross-sectional and longitudinal components, the personal identification number is composed of the household ID and the person number.

2. In the longitudinal component, the person number for any new person in the household shall be constructed by adding 1 to the highest person number used, for all the years of the survey and the household ID.

3. For the longitudinal component, the household ID shall be the household identification number of the household where the person first appears in the panel.

4. In the longitudinal component, the personal ID shall remain fixed over the length of time the person is in the survey.

3. INFORMATION TO BE COLLECTED

1. In the initial household, the full information required for current household members, the basic information for former household members and the basic information on household members in the previous wave that are no longer household members shall be collected.

2. In the split-off household, only the full information required for current household members shall be collected.

3. The full information required for current household members, the basic information for former household members and the basic information on household members in the previous wave that are no longer household members is set out in Commission Regulation (EC) No 1983/2003 [1]in the list of target primary variables.

4. Where a sample person is in the survey for more than one year, information shall be obtained on whether that person remained at the same address or moved to a different address from one year to the next.

4. RESPONDENT TO THE HOUSEHOLD QUESTIONNAIRE

1. Interview information relating to the whole household and general information on each household member shall be provided by (one or more) members of the household. The household respondent shall be chosen according to the following priorities:

— Priority 1: the person responsible for the accommodation.

— Priority 2: a household member aged 16 or over, who is best placed to give the information.

2. For the second and following waves of the longitudinal component of EU-SILC, the household respondent shall be chosen according to the following list of priorities:

— Priority 1: the household respondent in the last wave.

— Priority 2: a 'sample person' aged 16 or over, giving priority to the person responsible for the accommodation or best placed to give the information.

— Priority 3: a 'non-sample person' aged 16 or over.

5. SAMPLE IMPLEMENTATION

1. Each Member State shall follow appropriate procedures to maximise the response rates achieved, in accordance with its own 'best practices'. In the case of an interview survey, at least three call-backs shall be made before a household or individual is accepted as non-responding, unless there are conclusive reasons (such as a definite refusal to cooperate, circumstances endangering the safety of the interviewer, etc.) why this cannot be done.

2. EU-SILC shall be based on the selection and implementation of a probability sample. As a rule, the units enumerated in the survey shall be exactly the same units as those selected for the purpose in accordance with the sampling design, i.e. not substituted for by other units.

3. Records shall be kept of the outcome of each household and person selected for the sample, distinguishing (i) eligible units successfully enumerated, (ii) eligible units not successfully enumerated, (iii) selected units not found to be eligible for the survey and (iv) units not successfully enumerated where the eligibility status could not be determined. Reasons for the failure to enumerate successfully eligible units shall be recorded, identifying at least the main categories such as non-contact, refusal and incapacity to respond (such as being ill). Circumstances of units found to be ineligible or in the uncertain eligibility category shall also be recorded.

4. The proxy rate shall be kept as limited as possible for the income personal variables and for any variables required for at least one household member aged 16 or more.

5. In the cross-sectional component and in the first year of each panel constituting the longitudinal component, controlled substitutions may be allowed only where the response rate falls below 60 % and one of the following situations arises:

— address non-contacted because it cannot be located or is inaccessible,

— address contacted but interview not completed because the household refuses to cooperate, the entire household is temporarily away or the household is unable to respond.

[1] See page 34 of this Official Journal.

No substitutions shall be allowed for second and subsequent waves of a panel in the longitudinal component.

Procedures shall be followed to ensure that the process of substitution is controlled to the maximum extent possible. These include the use of a design which ensures that substitutes are selected so as to match the substituted units closely for important characteristics. The substitutions shall be allowed only after special efforts have been made to convince refusals, and properly scheduled call-backs have been made to other non-respondents. A household shall not be substituted by another non-sample household at the same address.

6. SURVEY DURATION AND TIMING

1. The interval between the end of the income reference period and the time of the interview for the respondent concerned shall be limited to eight months as far as possible. This applies both to the household and personal samples, and irrespective of whether the reference period used is fixed in terms of calendar dates for the whole sample or is a moving reference period determined according to the timing of the interview for the household or person concerned.

2. By way of exception to paragraph 1, if the income variables are collected from registers the interval between the end of the income reference period and the time of interview for current variables shall be limited to 12 months.

3. Where all the data are collected through field interviewing and a fixed income reference period is used, the total duration of the data collection of the sample shall be limited to four months as far as possible.

4. Where the data are collected through field interviewing using a moving income reference period and the fieldwork duration exceeds three months, the total annual sample shall be shared approximately equally between the fieldwork months. In this case, the total fieldwork duration for the cross-sectional component and each wave of the longitudinal component shall not exceed 12 months.

5. For the longitudinal component, the collection or compilation of data, for a given unit (household or person), between successive waves shall be kept as close as possible to 12 months.

7. IMPUTATION PROCEDURES

7.1. **Kinds of missing data**

1. Coverage and sample selection errors: for instance, where units in the target population are not represented in the sampling frame, or where the unit selection probabilities are distorted, or other sample selection errors occur.

2. Unit non-response: refers to absence of information on whole units (households and/or persons) selected into the sample.

3. Partial unit non-response: describes the situation where some but not all household members eligible and selected for the individual interview have been successfully enumerated.

4. Item non-response: refers to the situation where a sample unit has been successfully enumerated, but not all the required information has been obtained.

7.2. **Characteristics of an imputation procedure**

Two types of approaches (which may be used in combination) could be applied to the data:

— Imputation: refers to the generation of missing information on the basis of statistical relationships internal to the data set.

— Modelling: makes use of substantive relationships and information external to the data set.

The procedure applied to the data should preserve variation of and correlation between variables. Methods that incorporate 'error components' into the imputed values shall be preferable to those that simply impute a predicted value.

Methods which take into account the correlation structure (or other characteristics of the joint distribution of the variables) shall be preferable to the marginal or univariate approach.